Simulování náhodných veličin

Proč potřebujeme simulovat náhodné veličiny ?

- Pravděpodobnost chování řady statistických postupů (odhadů, testových statistik, procedur, apod.) nelze studovat exaktně.
- Rozdělení může být příliš složité nebo prakticky nepoužitelné.
- Sledujeme chování při konečných rozsazích výběrů pomocí simulací.
- Nechť X je náhodná veličina a $X_1, X_2, \ldots, X_n$ je náhodný výběr, pak $\mu_n = \frac{1}{n}\sum_{i=1}^n g(X_i) \overset{n\to\infty}{\Longrightarrow} \mu = \mathrm{E}\, g(X)$ (s pravděpodobností 1), speciálně
  - $\overline{X} = \frac{1}{n}\sum_{i=1}^n X_i \overset{n\to\infty}{\Longrightarrow} \mathrm{E}\, X$
  - $\frac{1}{n}\sum_{i=1}^n (X_i - \overline{X})^2 \overset{n\to\infty}{\Longrightarrow} \mathrm{var}\, X$
  - $\frac{1}{n}\,|X_n \leq c| \overset{n\to\infty}{\Longrightarrow} P(X \leq c)$
  - $\mathrm{median}(X_1, X_2, \ldots, X_n) \overset{n\to\infty}{\Longrightarrow} \mathrm{median}(X)$
- Je-li proces generovaný pomocí algoritmu, **není náhodný**. Důležité je, aby se proces **choval jako náhodný**.

Generování náhodných (pseudonáhodných) čísel z rovnoměrného rozdělení

- lineární kongruentní generátor $x_i = (ax_{i-1} + c) \mod m \, (i = 1, 2, \dots)$
- multiplikativní rekurzivní generátor $x_i = (a_1 x_{t-1} + \dots + a_k x_{t-k}) \mod m \, (i = k, k+1, \dots)$
- kombinované generátory, např. $x_i = (1403580 x_{i-2} - 810728 x_{i-3})$ mod $m_1 = 2^{32} - 209$ a $y_i = (527612 y_{i-2} - 1370589 y_{i-3})$ mod $m_2 = 2^{32} - 22853$ a $u_i = \frac{x_i - y_i + m_1}{m_1 + 1}$ pokud $x_i \leq y_i$ a $u_i = \frac{x_i - y_i}{m_1 + 1}$ pokud $x_i > y_i$
- kvalitu pseudonáhodných generátorů testuje řada testů
- Více např. Dirk P. Kroese, Thomas Taimre, Zdravko I. Botev: Handbook of Monte Carlo Methods,2011. ISBN: 9780470177938

Obecné metody simulování náhodných veličin

- metoda inverzní transformace

  využívá znalost distribuční funkce
- zamítací metoda

  využívá znalost funkce hustoty
- doplňková přijímací metoda
- kompoziční metoda

  lze využít pokud umíme napsat funkci hustoty jako součet jiných funkcí hustoty nebo jako transformace jiných funkcí hustoty

Metoda inverzní transformace

Uvažujeme náhodnou veličinu X s distribuční funkcí $F(x)$ a inverzní distribuční funkcí $F^{-1}(t) = \inf\{x : F(x) \geq t\}$, pak platí: pokud U má rovnoměrné rozdělení na intervalu $0, 1$, pak X má rozdělení s distribuční funkcí $X = F^{-1}(U)$.

Pro generování výběru $x_1, x_2, x_3, \ldots$ lze použít následující postup

- Nageneruji u_i z rovnoměrného rozdělení na intervalu $0, 1$
- $x_i = F^{-1}(u_i)$

Metoda je vhodná pro situace, kdy umím vyjádřit explicitně fci F^{-1}. Funkci F^{-1} lze v některých případech nahradit vhodnou aproximací. Hodnotu x_i lze získat i numerickým řešením rovnice $F(x_i) = u_i$.

Zamítací metoda

Nechť X je náhodná veličina s hustotou $f(x)$ a nechť existuje dominující funkcí hustoty $g(x)$ a konstanta $c \geq 1$ tak, že $f(x) \leq c \cdot g(x)$, pak lze pro generování výběru $x_1, x_2, x_3, \ldots$ použít následující postup

- nageneruji x_i s hustotou $g(x)$
- nageneruji v_i z z rovnoměrného rozdělení na intervalu $0, 1$ nezávislou na X
- vypočtu $T_i = c \cdot \frac{g(x_i)}{f(x_i)}$
- pokud platí $v_i \cdot T_i \leq 1$ je x_i akceptováno do výběru

Zamítací metoda pro funkci hustoty s omezeným nosičem

Pokud funkce $f(x)$ je nenulová na konečném intervalu a, b (označme $M = \max_{a \leq x \leq b} f(x)$), pak lze použít následující postup

- nageneruji u_i a v_i dvě nezávislé veličiny z $R(0,1)$
- $x_i = a + (b - a) \cdot u_i$
- pokud platí $v_i \cdot M \leq f(x_i)$ je x_i akceptováno do výběru

Výhoda: není třeba znát distribuční funkci a inverzní funkci F^{-1}
Nevýhoda: pokud funkce $f(x)$ vymezuje v obdélníku $(a, b) \times (0, M)$ malou plochu, bude zamítnuto mnoho dvojic u_i, v_i
Modifikací je metoda o sevření.

Generování náhodných vektorů pomocí inverzní metody

Uvažujme dvourozměrnou náhodnou veličinu s distribuční funkcí $F(x, y)$. Označme $F(x)$ marginální distribuční funkci a $F(y|x)$ podmíněnou distribuční funkci.

Pro generování výběru (x_i, y_i) použijeme následující postup

- nageneruji u_i a v_i dvě nezávislé veličiny z $R(0, 1)$
- $x_i = F^{-1}(u_i)$
- $y_i = F^{-1}(v_i|x_i)$

Generování náhodných vektorů pomocí zamítací metody

Uvažujme dvourozměrnou náhodnou veličinu s funkcí hustoty $f(x, y)$.
Předpokládejme, že funkce je nenulová na omezené oblasti A. Označme maximum $M = \max_{(x,y)\in A} f(x, y))$.
Pro generování výběru (x_i, y_i) použijeme následující postup

- nageneruji trojici x_i, y_i, w_i rovnoměrně z oblasti vymezené plochou A a výškou M
- pokud platí $w_i \leq f(x_i, y_i)$ je (x_i, y_i) akceptováno do výběru